

UC Irvine

UC Irvine Previously Published Works

Title

Rainfall frequency analysis for ungauged sites using satellite precipitation products

Permalink

<https://escholarship.org/uc/item/9721f52t>

Authors

Gado, TA

Hsu, K

Sorooshian, S

Publication Date

2017-11-01

DOI

10.1016/j.jhydrol.2017.09.043

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



Research papers

Rainfall frequency analysis for ungauged sites using satellite precipitation products

Tamer A. Gado^{a,*}, Kuolin Hsu^b, Soroosh Sorooshian^b^a Department of Irrigation and Hydraulics Engineering, Faculty of Engineering, Tanta University, Tanta, Egypt^b Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California Irvine, E/4130 Engineering Gateway, Irvine, CA 92697-2175, USA

ARTICLE INFO

Article history:

Received 10 April 2017

Received in revised form 22 September 2017

Accepted 24 September 2017

Available online 28 September 2017

Keywords:

Regional rainfall frequency analysis

Extreme rainfall

Satellite precipitation

PERSIANN

Probability matching method

Index flood

Ungauged site

ABSTRACT

The occurrence of extreme rainfall events and their impacts on hydrologic systems and society are critical considerations in the design and management of a large number of water resources projects. As precipitation records are often limited or unavailable at many sites, it is essential to develop better methods for regional estimation of extreme rainfall at these partially-gauged or ungauged sites. In this study, an innovative method for regional rainfall frequency analysis for ungauged sites is presented. The new method (hereafter, this is called the RRFA-S) is based on corrected annual maximum series obtained from a satellite precipitation product (e.g., PERSIANN-CDR). The probability matching method (PMM) is used here for bias correction to match the CDF of satellite-based precipitation data with the gauged data. The RRFA-S method was assessed through a comparative study with the traditional index flood method using the available annual maximum series of daily rainfall in two different regions in USA (11 sites in Colorado and 18 sites in California). The leave-one-out cross-validation technique was used to represent the ungauged site condition. Results of this numerical application have found that the quantile estimates obtained from the new approach are more accurate and more robust than those given by the traditional index flood method.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

An accurate estimation of extreme rainfall (magnitude, duration, and frequency) is fundamental for the planning and design of various hydraulic structures. Therefore, many studies have focused on the development of methods for improving the accuracy of extreme rainfall estimation. In this regard, rainfall frequency analysis (RFA) is commonly used to estimate the rainfall rate/volume for a given return period at a given site of interest. As precipitation records are often unavailable at many sites, regional rainfall frequency analysis (RRFA) has become an essential tool for extreme rainfall estimation for these ungauged sites. In addition, RRFA is often used to improve the accuracy of at-site extreme rainfall estimation at gauged sites. For example, the RRFA was used to analyze the extraordinary storm that occurred in the City of Fort Collins, Colorado, in July 1997 (Sveinsson et al., 2002). The study concluded that the city's storm drainage design criteria were

underestimated. In general, RRFA consists of two main steps: the identification of groups of hydrologically homogeneous regions, and the application of a regional estimation method within each delineated homogeneous group (Gado and Nguyen, 2016).

Determination of homogeneous regions is the first step in RRFA. The goal of this step is to combine the extreme rainfall information from a group of similar sites for improving the estimation of rainfall at any site in that group. Traditional techniques of delineating homogeneous groups of sites are based on their geographical locations and/or administrative and political boundaries. Recent techniques, such as cluster analysis (Tasker, 1982), discriminant analysis (Wiltshire, 1986), discordancy measure (Hosking and Wallis, 1993), region of influence (ROI) (Burn, 1990), and canonical correlation analysis (Cavadias et al., 2001) have been recommended for homogeneous region delineation. The second step in RRFA is the application of a regional extreme rainfall estimation method within each delineated homogeneous group. Regional extreme rainfall estimation can be carried out by the index flood method (Dalrymple, 1960) or the regional regression method (Benson, 1962). Although the index flood method has been recommended for extreme hydrologic event estimation for ungauged sites, the accuracy of extreme rainfall estimates for such sites based on both techniques is quite limited.

* Corresponding author.

E-mail addresses: tamer.gado@f-eng.tanta.edu.eg, tamer.gado@mail.mcgill.ca, tamergado@hotmail.com (T.A. Gado), kuolinh@uci.edu (K. Hsu), soroosh@uci.edu (S. Sorooshian).

Recently, several products of satellite-based precipitation data have been developed in order to estimate precipitation data with high spatial and temporal resolution and near-global coverage. These precipitation data are extremely valuable in regions where ground measurements are very sparse or even nonexistent. However, the quality of precipitation estimation from all satellite products must be evaluated over all climatic and geographic regions of the world. The justification of satellite-based precipitation data with regard to gauged data has been investigated by numerous studies in variety of temporal and spatial scales in different regions in the world (e.g., Yilmaz et al., 2005; Chiang et al., 2007; Boushaki et al., 2009; Jiang et al., 2010; Gourley et al., 2011; AghaKouchak et al., 2011; Behrangi et al., 2011; Katiraei-Boroujerdy et al., 2013; Xue et al., 2013; Jamandre and Narisma, 2013; Wasko et al., 2016). Lately, few studies have used remote sensing data for rainfall frequency analysis and for the derivation of IDF curves in different regions (e.g., Endreny and Imbeah, 2009; Awadallah et al., 2011; Wright et al., 2013; Eldardiry et al., 2015; Marra et al., 2017).

The current development of satellite-based precipitation retrieval algorithms makes them prominent for extreme rainfall estimation at ungauged sites. Therefore, satellite-based precipitation estimation can be used as a potential data source for improving the accuracy of extreme rainfall estimation at ungauged sites. Recent developed satellite measurements at high spatial and temporal resolution have high potential for being used for hydrologic applications of remote regions (e.g., Yilmaz et al., 2005; Chiang et al., 2007; Su et al., 2008; Ashouri et al., 2015). Although several near real-time precipitation retrieval algorithms were developed recently (Joyce et al., 2004; Huffman et al., 2007, 2015; Hsu et al., 1997; Sorooshian, 2000; Hong et al., 2004; Ushio et al., 2009; Kuligowski, 2002), for studying climate extremes, long-term historical data at high spatial and temporal resolutions are required. Satellite-based estimates for climate study are mostly processed at lower spatial/temporal resolutions. GPCP (Huffman et al., 1997) and CMAP (Xie and Arkin, 1997) products, for example, are made available at monthly scale. GPCP 1DD (Huffman et al., 2007) and TRMM 3B42 (V7) data are processed at high spatial resolution with data made available since 1998 to near current time. One new climatic data product, named PERSIANN-Climate Data Record (PERSIANN-CDR) is a daily 0.25° precipitation product that covers the area between 60°S and 60°N latitude and 0° and 360° longitudes. The data covers from 1 January 1983 to near-current time (Ashouri et al., 2015). PERSIANN-CDR provides a unique data source for long-term (30+ years) hydroclimate analysis.

Scanning the literature for satellite-based precipitation analysis reveals that little has been done in regard to evaluation of satellite-based precipitation products using rainfall frequency analysis in order to estimate extreme rainfall at ungauged sites (or enhance extreme rainfall estimation at gauged sites). Hence, in order to improve the accuracy of extreme rainfall estimation at ungauged sites, this research will aim at two main objectives: (1) to assess the at-site extreme rainfall estimation by applying the rainfall frequency analysis using historical data from a satellite based precipitation product (PERSIANN-CDR); and (2) to develop a new method for regional rainfall frequency analysis using satellite data; hereafter, the new method is called the RRFA-S. The new method is based on corrected annual maximum series of rainfall data from PERSIANN-CDR using the probability matching method (PMM). Finally, the new method was assessed through a comparative study with the traditional index flood method using the available annual maximum series of daily rainfall in two different regions in USA (11 sites in Colorado and 18 sites in California). More specifically, the estimation of extreme rainfall at ungauged sites using satellite based precipitation data (the new method, RRFA-S) was compared with that using gauged precipitation data (the index flood method, IFM).

2. Methodology

2.1. Data

To illustrate the application of the proposed approach, a case study was carried out using the annual maximum daily precipitation series from two regions in the USA: 11 stations in Colorado and 18 stations in California (Fig. 2). These two regions were already investigated as homogenous regions in these previous studies: Sveinsson et al. (2002) and Bonnin et al. (2011). The database used here was extracted from the National Oceanic and Atmospheric Administration's (NOAA) National Weather Service (NWS) http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_series.html (Bonnin et al., 2011). Table 1 provides some of the characteristics of the selected stations used in this study. The length of the rainfall series varies from 20 to 107 years. The selection of the stations is made in such a way that the observed series are stationary in the mean. Here, the stationarity of the series was assessed by the Mann-Kendall test.

For acquiring good global temporal and spatial precipitation data, there have been some high-resolution global satellite precipitation products operationally available, including precipitation estimation from remotely sensed information using artificial neural networks (PERSIANN), the PERSIANN-Climate Data Record (PERSIANN-CDR) (Ashouri et al., 2015; Hsu et al., 2014). The purpose of the development of the PERSIANN-CDR precipitation product is to study the spatial and temporal characteristics of precipitation in a scale relevant to climate studies. PERSIANN Precipitation Climate Data Record (PERSIANN-CDR) is generated from PERSIANN algorithm using GridSat-B1 infrared data and it is adjusted using GPCP monthly product at 2.5° monthly. The PERSIANN-CDR is a daily 0.25° precipitation product that covers the area between 60°S and 60°N latitude and 0° and 360° longitude and spans the period 1 January 1983 to near-current time (Ashouri et al., 2015; Hsu et al., 2014). The satellite precipitation data was obtained from the PERSIANN-CDR through NCDC CDR Program: <http://www.ncdc.noaa.gov/cdr/index.html> (Sorooshian et al., 2014).

The annual maximum daily series should result from high resolution time series, e.g., sliding a 24 h window and then extracting the annual maximum value (Papalexioiu et al., 2016). However, in this research, the annual maximum series was extracted as the maximum of daily observations in the year because the PERSIANN-CDR is a daily precipitation product.

2.2. Index flood method (IFM)

The index flood method starts by standardizing the extreme rainfall information at each site in the region by dividing the annual maximum rainfall data by a scale or an index (typically, the mean of annual maximum series, μ). A regional distribution is then applied on the pooled standardized extreme rainfall information from the stations in the region to get the standardized extreme rainfall estimate corresponding to the desired return period. Finally, the regional extreme rainfall estimates for the site of interest are calculated by multiplying its index and the standardized extreme rainfall estimates. In case of ungauged sites, different techniques (see below) can be used to estimate the index for the site of interest. In this study, the index is the at-site mean extreme rainfall.

An ungauged site is a site where no data is available. The index (i.e., μ) at an ungauged site can be estimated using available data at gauged sites within the homogeneous region that includes that ungauged site. In the RRFA, some techniques have been used to estimate the index for ungauged sites. In this study, three techniques were used as follows:

Table 1

Summary of the characteristics of the selected stations in Colorado and California.

Station ID	Station Name	Latitude	Longitude	Elevation (m)	Sample Size (year)	Sample Mean (mm)
Colorado						
USC00050834	BONNY DAM 2 NE	39.61667	−102.18333	1133	67	49
USC00050304	ARAPAHOE	38.85	−102.18333	1223	67	47
USC00056299	PAOLI	40.6167	−102.4667	1190	23	45
USC00052535	ECKLEY	40.1089	−102.4867	1193	52	45
USC00054380	JOES	39.65	−102.68333	1296	66	46
USC00052446	EADS	38.48333	−102.78333	1299	107	48
USC00057519	SEIBERT	39.11667	−102.86667	1445	59	48
USC00050109	AKRON 4 E	40.15	−103.15	1384	99	45
USC00054172	HUGO 1NW	39.13333	−103.46667	1532	74	37
USC00056136	ORDWAY 21N	38.53	−103.7058	1453	64	41
USC00055922	NEW RAYMER	40.6089	−103.8461	1458	56	44
California						
USC00040244	APPLE VALLEY	34.5167	−117.217	894	29	24
USC00040418	BACKUS RANCH	34.95	−118.183	808	28	26
USC00040519	BARSTOW	34.9	−117.033	659	56	22
USC00040521	BARSTOW FIRE STATION	34.8878	−117.023	707	76	23
USC00041488	CANTIL 2 E	35.3167	−117.933	598	20	23
USC00042257	DAGGETT FCWOS	34.8536	−116.786	584	55	23
USC00042771	EL MIRAGE FIELD	34.5897	−117.631	899	40	26
USC00044747	LANCASTER	34.6833	−118.117	732	30	27
USC00044749	LANCASTER FSS	34.7333	−118.217	713	28	37
USC00045002	LLANO EBERLE RANCH	34.4667	−117.75	1165	21	34
USC00045182	LUCERNE VALLEY 2 W	34.45	−116.983	909	25	20
USC00045756	MOJAVE	35.0492	−118.162	834	87	28
USC00046122	NEENACH	34.8	−118.583	881	29	38
USC00046624	PALMDALE	34.5833	−118.1	791	98	32
USC00047253	RANDBURG	35.3692	−117.653	1088	79	31
USC00048826	TEHACHAPI	35.1333	−118.45	1224	80	34
USC00049325	VICTORVILLE PUMP PLANT	34.535	−117.306	871	80	27
USC00049345	VINCENT FS FC 120	34.4883	−118.141	956	74	37

1. The regional index method: Annual maximum series (AMS) of extreme rainfall data at all gauged sites in a homogeneous region are used to compute the index for each site. The regional index (M) is computed using the AMS lumped from data of all gauged sites within the homogeneous region.

$$M = \frac{\sum_{i=1}^n \left(\frac{\sum_{j=1}^m X_{ji}}{m} \right)}{n} \quad (1)$$

where X_{ji} is the j th record of data at the i th gauged site, m is the number of records at that gauged site, and n is the number of gauged sites in the homogeneous region. The estimated regional index is considered as the same for any site in the homogeneous region including the ungauged site (i.e., $\mu = M$).

2. The nearest gauged site: The index for the ungauged site is estimated equal to that of the nearest gauged site.
3. The inverse squared distance method (ISDM): The index (μ) at the ungauged site is estimated from those at the gauged sites in a homogenous region by using the ISDM:

$$\mu = \sum_{i=1}^n W_i \mu_i \quad (2)$$

where W_i is the weight assigned to the i^{th} gauged site and is given by:

$$W_i = \frac{d_i^{-2}}{\sum_{i=1}^n (d_i^{-2})} \quad (3)$$

where d_i is the distance from the ungauged site to the i^{th} gauged site.

Application of the generalized extreme value (GEV) distribution to model the extreme hydrologic event series has been advocated by several researchers (e.g., Hosking et al., 1985; Lettenmaier and Potter, 1985). The cumulative distribution function (cdf) $[F(q)]$ for the GEV distribution is given as:

$$F(p) = \exp \left[- \left(1 - \frac{\kappa(p - \xi)}{\alpha} \right)^{1/\kappa} \right] \quad \kappa \neq 0 \quad (4)$$

$$= \exp \left[- \exp \left(- \frac{p - \xi}{\alpha} \right) \right] \quad \kappa = 0$$

where: p is the observation extreme precipitation; and ξ , α , and κ are respectively the location, scale, and shape parameters of the distribution.

The index flood frequency estimation procedure that employs the GEV distribution with L-moments parameter estimation (GEV/L) has been found to be an efficient way of combining hydrologic (e.g., rainfall) data in regional frequency analysis (e.g., National Environmental Research Council, 1975; GREHYS, 1996; Hosking and Wallis, 1997).

2.3. Probability matching method (PMM)

The probability matching method (PMM), initially proposed by Calheiros and Zawadzki (1987), is an attractive method for correcting the remote sensed rainfall data in regard to gauged data, because the frequency distributions of two datasets are matched percentile by percentile ensuring that the extremes are matched. Here, in the PMM, it is assumed that the satellite precipitation data has the same probability of occurrence as the gauged rainfall data. The PMM is based on matching the cumulative density functions (CDF) of both data types as described in Eq. (5) and shown in Fig. 1.

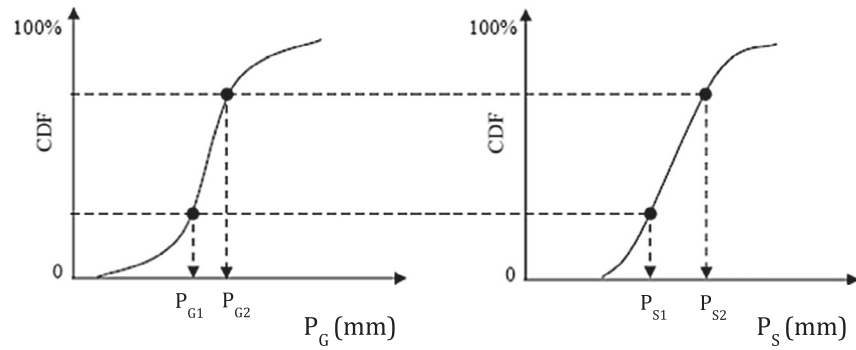


Fig. 1. The probability matching method (PMM) (modified after Piman et al., 2007).

Table 2

Values of the GEV parameters and the 100-year extreme rainfall for the three data types at stations Eckley, CO and Randsburg, CA.

Station	Data Type	GEV Parameters			P_{100} (mm)
		Location (ξ)	Scale (α)	Shape (κ)	
Eckley, CO	GAMS	36.80	12.87	−0.065	106
	SAMS	21.69	6.91	−0.096	62
	CSAMS	36.60	12.92	−0.072	107
Randsburg, CA	GAMS	23.65	14.06	0.023	85
	SAMS	19.44	9.01	−0.043	65
	CSAMS	23.54	14.00	0.023	85

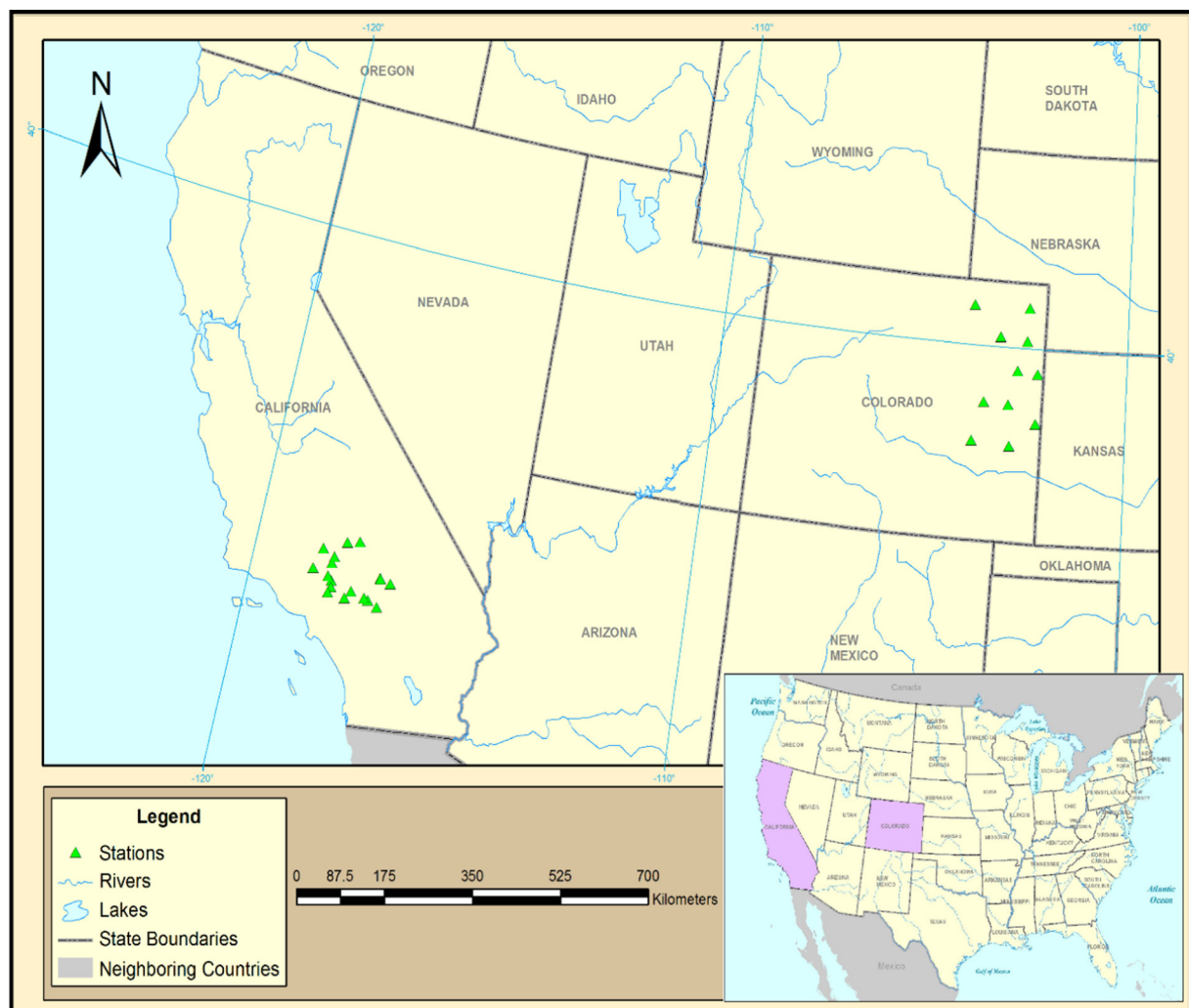


Fig. 2. Geographical locations of the selected gauging stations in Colorado and California.

$$\int_{P_G}^{\infty} P(P_G) dP_G = \int_{P_S}^{\infty} P(P_S) dP_S \quad (5)$$

where $P(P_G)$ and $P(P_S)$ are the probability density functions of gauged and satellite precipitation data, respectively. P_G and P_S having the same CDF values are matched as pairs and then these pairs are used to determine the relationship between the satellite precipitation and the corrected one. The PMM was used here to avoid sampling volume and timing problems in sampling data comparison (Piman et al., 2007).

2.4. Regional rainfall frequency analysis using satellite precipitation data (RRFA-S)

The main purpose of the new method is to estimate extreme rainfall quantiles at an ungauged site by using data from a satellite precipitation product at that site and gauged data from neighborhood sites in a homogeneous region. The proposed method consists of:

1. Extracting AMS from gauged data for every site in a homogeneous region (gauged AMS will be called GAMS);

2. Extracting AMS from a satellite precipitation product (e.g., PERSIANN-CDR) for every site in a homogeneous region (satellite AMS will be called SAMS);
3. Correcting SAMS, obtained from step 2, by using the PMM in order to match the CDFs of both SAMS and GAMS (obtained from step 1) for every site (corrected SAMS will be called CSAMS);
4. Establishing a relationship between CSAMS and SAMS for all sites in the homogeneous region;
5. Estimating the CSAMS for any ungauged site in the homogeneous region by the modelled relationship in step 4, knowing the SAMS for that site; and
6. Estimating extreme rainfall quantiles by applying the RFA (e.g., GEV distribution with L-moments) using the CSAMS at the ungauged site.

2.5. Comparison of extreme rainfall estimation methods at ungauged sites

As mentioned previously, the comparison of the two approaches (RRFA-S and IFM) for the estimation of extreme rainfall at ungauged sites was carried out in this study. In both approaches, the leave-one-out cross-validation procedure provides a suitable

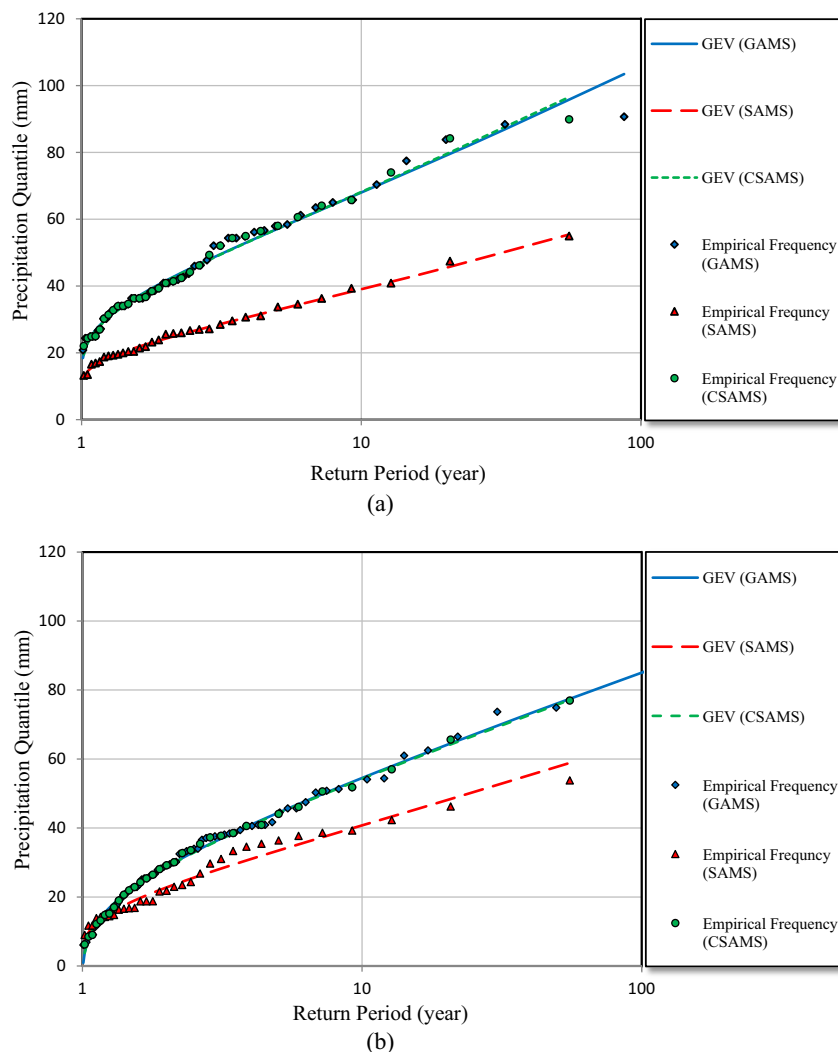


Fig. 3. Precipitation quantiles of different return periods for the three data types (GAMS, SAMS, and CSAMS), (a) Station Eckley, Colorado, (b) Station Randsburg, California.

tool to simulate the ungauged site condition. That is, one site is removed from the data base and the model is developed using the data from the remaining sites. The model is in turn used to predict the quantiles for the site not used in the model development. The process is repeated until every site is removed once. For the new method (RRFA-S), the predicted quantiles at every site will be obtained from the fitted distribution (GEV/L) using the CSAMS, the corrected annual maximum series obtained from the satellite precipitation product (PERSIANN-CDR) by using the PMM. In both approaches, the predicted quantiles are compared with those obtained from the at-site fitted distribution (GEV/L) using the observed (gauged) rainfall data. The two approaches are compared according to their ability to provide reliable estimates of extreme rainfall quantiles (i.e., close to at-site fitted quantiles) for ungauged sites. Here, seven values of return periods were considered (5, 10, 25, 50, 100, 200, and 1000 years).

To assess the performance of both extreme rainfall estimation procedures, the at-site fitted GEV/L distribution is used to compute the quantiles for the considered return periods. These quantiles are obtained from the fitted GEV/L distribution using the available at-site data. Therefore, they could be considered as representative values of the unknown true rainfall quantiles for the specific site. Then, in this study, an extreme rainfall estimation procedure is considered accurate if it can provide extreme rainfall estimates that are close to the fitted at-site values for the site studied. An objective assessment of the performance of both methods can be obtained using the following four numerical criteria:

1. Mean bias (BIAS)

$$\text{BIAS} = \frac{1}{M} \sum_{i=1}^m (P_{T,i} - \hat{P}_{T,i}) \quad (6)$$

2. Relative mean bias (BIAS_r)

$$\text{BIAS}_r = \frac{1}{M} \sum_{i=1}^m \left(\frac{P_{T,i} - \hat{P}_{T,i}}{P_{T,i}} \right) \quad (7)$$

3. Root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^m (P_{T,i} - \hat{P}_{T,i})^2} \quad (8)$$

4. Relative mean square error (RMSE_r)

$$\text{RMSE}_r = \sqrt{\frac{1}{M} \sum_{i=1}^m \left(\frac{P_{T,i} - \hat{P}_{T,i}}{P_{T,i}} \right)^2} \quad (9)$$

where $P_{T,i}$ and $\hat{P}_{T,i}$ are respectively the at-site (fitted) and regional (predicted) quantiles (corresponding to a given return period T) at station i ; and M is the number of total stations.

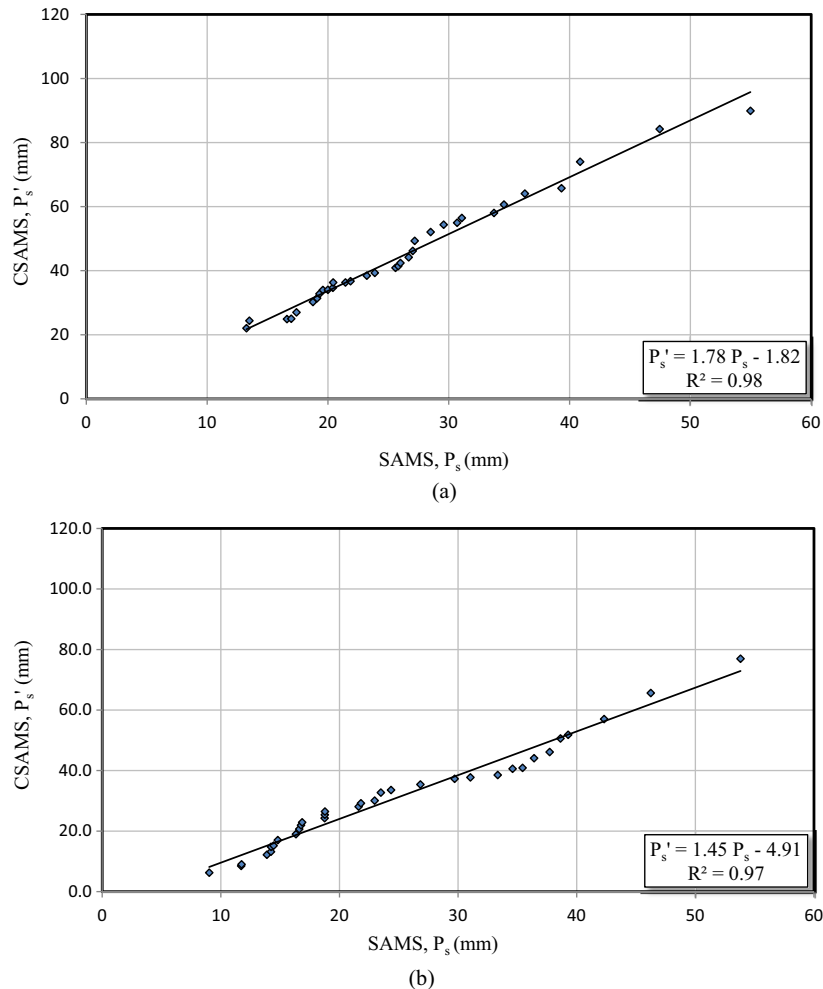


Fig. 4. Relationship between the AMS from the PERSIANN-CDR data (SAMS) and the corrected AMS (CSAMS) using the PMM, (a) Station Eckley, Colorado, (b) Station Randsburg, California.

3. Results

3.1. Estimation of quantiles for gauged sites

Using historical data from both gauged sites and the satellite precipitation product PERSIANN-CDR, at-site extreme rainfall quantiles were estimated by applying the GEV distribution along with the L-moments. Table 2 and Fig. 3 give an example of the application of GEV/L on two stations in two different states (Eckley, CO and Randsburg, CA). It can be shown from Fig. 3 that the GEV distribution fits well both gauged and satellite data for both stations. However, the extreme rainfall quantiles obtained from the satellite data are far from those obtained from the gauged data (Fig. 3 & Table 2).

Gauged data (GAMS) are point measurements while satellite estimation (SAMS) is an area measurement at 0.25° lat-long. Grid data consider an area average of high number of gauges in the grid coverage. Accordingly, SAMS are usually lower in absolute value than GAMS, i.e., the location parameter of the GEV distribution of the SAMS is lower than that of GAMS (Table 2). Moreover, there is a considerable variability of gauge precipitation measurement under the coverage of 0.25°. Because of “scale” difference from point to pixel, it is expected that gauge (point) measurements could have higher variability and extreme values than that of esti-

mates from grid-based precipitation at 0.25°. Consequently, results indicate that the scale parameter of the GEV distribution of the GAMS is much higher than that of SAMS (Table 2).

Furthermore, it can be concluded that the satellite precipitation data of Colorado stations has higher bias than that of California stations regarding extreme rainfall estimation (Fig. 3). That could be due to precipitation over Colorado having high spatial and temporal variability because of convective rainfalls and orographic rains in summer. In contrast, over California, the major rainfall season is the winter, when precipitation is mainly from large scale stratiform rain storms. This type of precipitation has a more homogeneous and wide spread distribution. Thus, the spatial variability is lower and rainfall estimates of 0.25° grid are close to point rain gauges over California.

Since the satellite precipitation data should have the same probability of occurrence as the gauged precipitation data, the probability matching method (PMM) is suitable for bias correction in this case. Accordingly, the PMM was used here in order to match the CDF of AMS of satellite data (SAMS) with that of gauged data (GAMS). Then, the resulted corrected satellite AMS (CSAMS) was used to estimate the corrected extreme rainfall quantiles by applying the GEV distribution. It can be concluded that the CSAMS give almost the same results as the GAMS for both stations (Table 2 & Fig. 3).

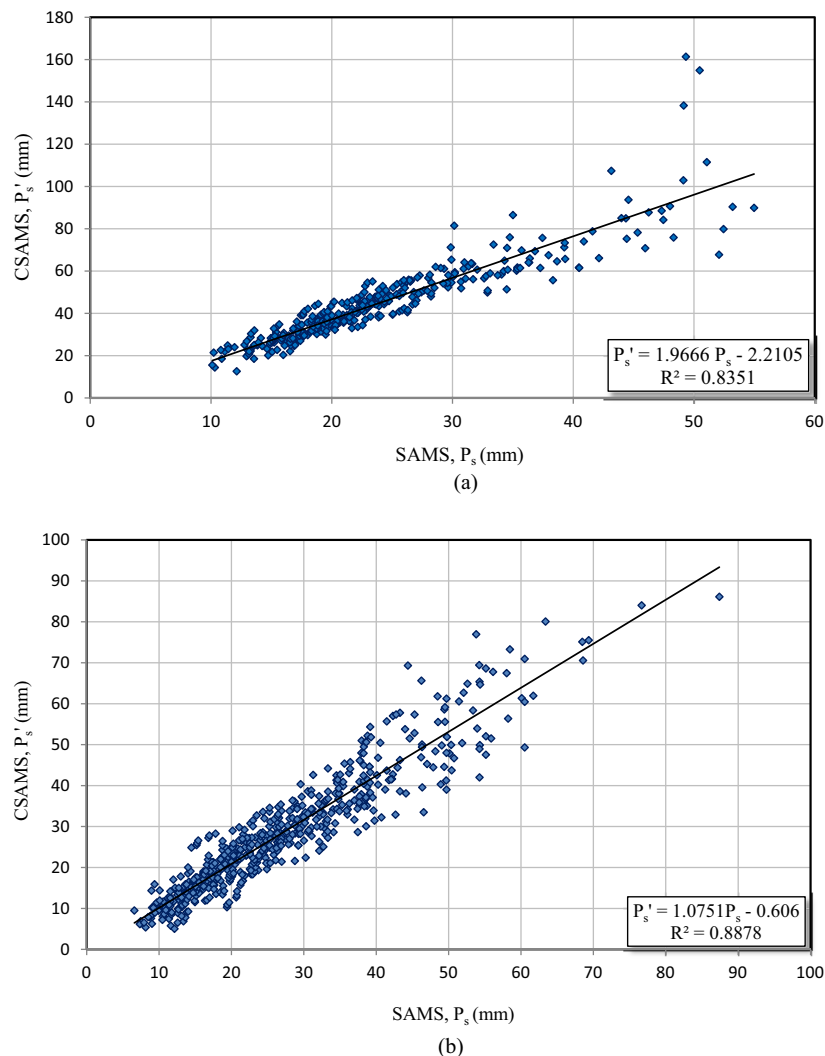


Fig. 5. Relationship between the SAMS and the CSAMS for the two regions, (a) Colorado, (b) California.

The relationships between the CSAMS and the SAMS at the two stations are shown in Fig. 4. It can be seen that the CSAMS has a strong linear relationship with the SAMS obtained from the PERSIANN-CDR at both sites.

3.2. Estimation of quantiles for ungauged sites

In the new method (RRFA-S), the relationship between the SAMS obtained from the PERSIANN-CDR and the corrected SAMS (CSAMS) using the PMM for all sites in the homogeneous region is drawn in Fig. 5 for both regions. From this linear relationship, one can estimate the CSAMS for any ungauged site in the region knowing the SAMS at that site. Then, the CSAMS can be used to estimate extreme rainfall quantiles by applying the rainfall frequency analysis (e.g., GEV distribution with L-moments).

The leave-one-out cross-validation procedure is used to compare the two approaches considered in this study (IFM and RRFA-S) in the two regions. The estimation procedure is applied for each regional estimation method and for each return period in both regions. As pointed out earlier, the GEV distribution, estimated

by the L-moments, is applied to calculate the at-site extreme rainfall estimation. For the index flood method, the estimation of the mean annual extreme rainfall (index flood) is required at an ungauged site, in order to estimate extreme rainfall quantiles at such site. Here, three techniques were used to obtain the required estimation of the index flood (see Section 2.1). A thoroughly comparative study of the three techniques was conducted in order to identify the best technique for accurately estimate the index flood at an ungauged site. The results (not included) indicated that the index flood estimates obtained from the inverse squared distance method (ISDM) are more accurate than those given by the other two methods. Therefore, the ISDM was selected here to estimate the index flood at ungauged sites.

The key idea in the suggested comparative methodology is that a regional extreme rainfall estimate (predicted) at a given site, obtained by neglecting all data at that site, can be compared with the at-site extreme rainfall estimate, given by the fitted GEV distribution using the observed data at the site of interest. Thus, the performance of the regional extreme rainfall estimation procedures is examined on the basis of the closeness of the predicted regional estimates with the fitted at-site values using graphical and numerical criteria. For purposes of illustration, the quantile-quantile (Q-Q) plot between the fitted and predicted extreme rainfall estimates for all studied return periods, based on the two methods, is shown in Fig. 6 for both regions. The Q-Q plot is a subjective graphical means of assessing the closeness of the predicted quantiles with the fitted ones. If the predicted and fitted quantiles are close to each other, then the points in the Q-Q plot should fall close to the 45° line. Fig. 6 shows that the majority of the predicted quantiles obtained from the proposed method are closer to the fitted quantiles than those given by the index flood method in both regions. Thus, the

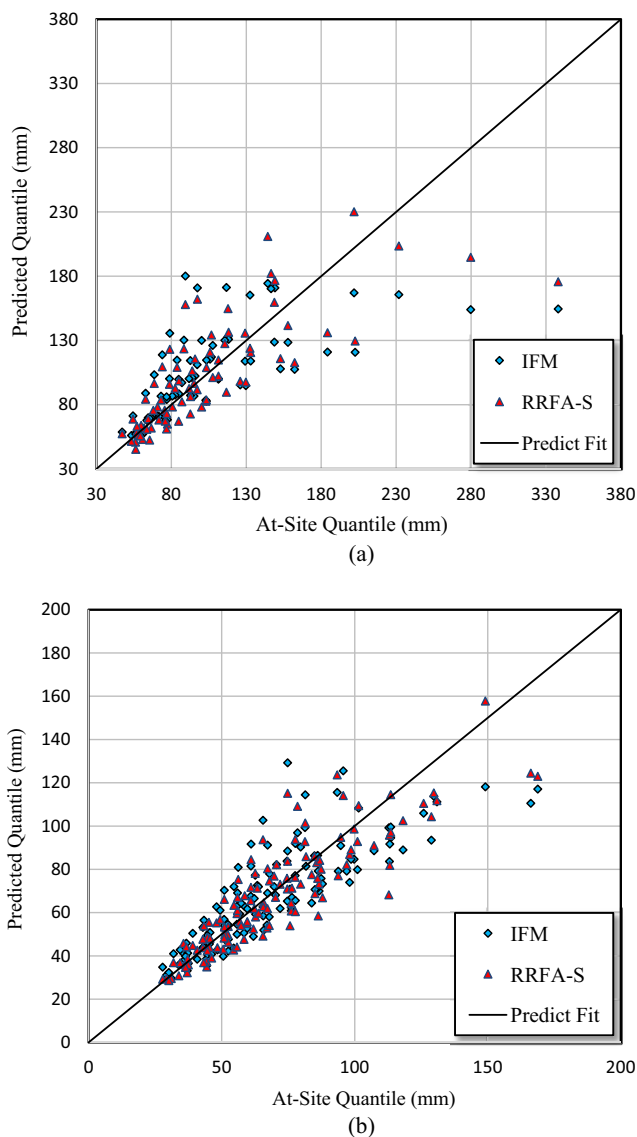


Fig. 6. Quantile-Quantile plots between at-site (fitted) and regional (predicted) estimated floods of all studied return periods for the two regions, (a) Colorado, (b) California.

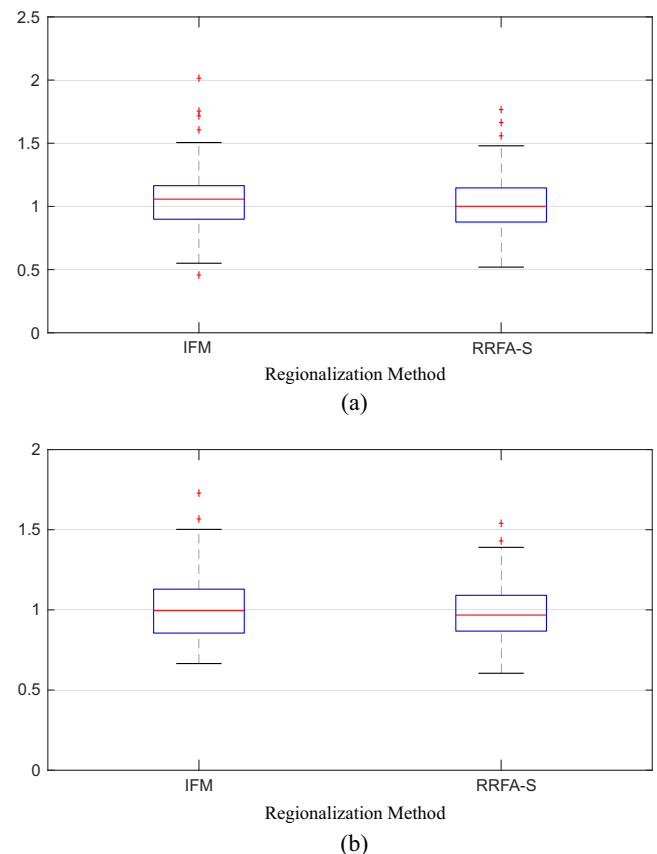


Fig. 7. Box plots of the ratios between regional and at-site quantiles of all studied return periods for the two regions, (a) Colorado, (b) California.

Table 3

Performance criteria of the two compared estimation methods (IFM and RRFA-S) for all considered return periods in the two regions.

Region	T (year)	BIAS (mm)		BIASr		RMSE (mm)		RMSEr	
		IFM	RRFA-S	IFM	RRFA-S	IFM	RRFA-S	IFM	RRFA-S
Colorado	5	−0.75	0.89	−0.02	0.01	4.09	5.95	0.08	0.11
	10	−1.15	1.08	−0.03	0.01	7.05	8.45	0.12	0.13
	25	−1.45	1.07	−0.04	0.00	13.42	13.38	0.17	0.17
	50	−1.34	0.96	−0.05	−0.02	20.36	18.73	0.22	0.20
	100	−0.75	0.85	−0.07	−0.03	29.49	25.85	0.28	0.24
	200	0.52	0.82	−0.09	−0.05	41.28	35.15	0.33	0.28
	1000	7.6	1.85	−0.14	−0.11	82.44	68.33	0.48	0.40
	All	0.38	0.44	−0.06	−0.01	37.86	19.32	0.27	0.18
California	5	0.60	0.33	−0.01	0.00	6.13	5.05	0.16	0.13
	10	0.72	0.73	−0.01	0.01	7.77	6.09	0.16	0.13
	25	0.92	1.32	−0.01	0.01	10.33	7.88	0.17	0.13
	50	1.13	1.83	−0.01	0.01	12.67	9.78	0.19	0.15
	100	1.42	2.40	−0.02	0.01	15.43	12.29	0.20	0.17
	200	1.80	3.04	−0.02	0.01	18.68	15.53	0.22	0.19
	1000	3.21	4.87	−0.03	0.00	28.51	26.24	0.27	0.25
	All	1.40	2.07	−0.01	0.01	15.88	13.63	0.20	0.17

Note: Bold values are the best values for each performance index and each return period among the values of the two methods to show the best method for each case.

new method is more accurate than the index flood method in regard to extreme rainfall estimates in the two regions.

To assess the robustness of the two methods, the box plots of the ratio of the regional (predicted) and at-site (fitted) quantiles for all studied return periods are drawn in Fig. 7 for both regions. It should be noted that if the predicted and observed quantiles are close to each other, then the ratio will be close to 1. In contrast, a ratio between the predicted and observed quantiles is far from 1 when the estimate is less accurate. Fig. 7 shows that the performance of both methods are almost equal.

For a more objective assessment, values of the performance criteria obtained for the two methods for all considered return periods are given in Table 3 for both regions. The optimum values (*i.e.*, smaller BIAS, BIASr, RMSE, and RMSEr values) for each performance criterion are highlighted in order to show the best method for each case. It can be seen that the quantile estimates obtained using the new approach (RRFA-S) are more accurate than those given by the index flood method (IFM) for most cases in both regions. In terms of BIASr, RMSE, and RMSEr, the overall improvements for the RRFA-S method, with respect of the IFM, are respectively 83%, 49%, and 33% for Colorado and 0%, 14%, and 15% for California (Table 3). In contrast, the IFM is better than the RRFA-S method in terms of BIAS by 14% (Colorado) and 32% (California). Notice that these percentages are calculated, in the case of all return periods, by dividing the difference between the values of the criterion of each methods by the biggest value. These results point to the advantage of using the RRFA-S approach over the IFM for extreme rainfall estimation at ungauged sites in the two studied cases.

4. Conclusions

In this study, a new regional estimation method of extreme rainfall at ungauged sites has been introduced. The new method involves applying regional rainfall frequency analysis using satellite precipitation data (RRFA-S). Specifically, in this method, a regional rainfall frequency technique (*e.g.*, GEV distribution with L-moments) can be applied on corrected annual maximum series (AMS) obtained from a satellite precipitation product (*e.g.*, PERSIANN-CDR). The probability matching method has been used here to correct the satellite based AMS. The data used in the new method (RRFA-S) combine satellite precipitation data at the target location, *i.e.*, an average rainfall data within the satellite pixel containing that location, with gauged data at neighborhood locations

forming a homogeneous region. The performance of this method has been investigated through a comparative study with the traditional index flood method. The comparative study, based on the leave-one-out cross-validation procedure, was conducted in two homogenous regions of two different states in USA (Colorado and California).

Results of the illustrative application have indicated that the RRFA-S method can provide more accurate extreme rainfall estimates than those given by the index flood method. The RRFA-S method takes the advantage of the development of satellite-based precipitation products which makes them prominent for extreme rainfall estimation at ungauged sites. Although the case study is based on hydrologic data from Colorado and California, the inferences made in this paper represent a starting point with respect to the analysis of regional extreme rainfall estimation using satellite based precipitation data. Thus, similar empirical studies should be carried out to assess the applicability of the new method to extreme rainfall estimation in other climatic regions.

In this study, the relationship between the satellite based AMS (SAMS) and the corrected satellite based AMS (CSAMS) is assumed to be fixed for homogeneous regions. Since this assumption implies the similarity of hydrologic conditions for all stations in a given homogeneous region, researchers should consider using this relationship in delineating such regions. Moreover, a recommendation for future studies is to evaluate different satellite rainfall products based on the accuracy of extreme rainfall estimation at ungauged sites by applying the new method in variety of temporal and spatial scales in different regions in the world.

References

- AghaKouchak, A., Behrangi, A., Sorooshian, S., Hsu, K., Amitai, E., 2011. Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *J. Geophys. Res.* 116, D02115. <https://doi.org/10.1029/2010JD014741>.
- Ashouri, H., Hsu, K., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: daily precipitation climate data record from multi-satellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* 96 (1), 69–83.
- Awadallah, A., ElGamal, M., ElMostafa, A., ElBadry, H., 2011. Developing intensity-duration-frequency curves in scarce data region: an approach using regional analysis and satellite data. *Engineering* 3, 215–226. <https://doi.org/10.4236/eng.2011.33025>.
- Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., Sorooshian, S., 2011. Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.* 397, 225–237.
- Benson, M.A., 1962. Factors Influencing the Occurrence of Floods in a Humid Region of Diverse Terrain. U.S Geological Survey, Washington, DC.

- Bonnin, G.M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2011. Precipitation-Frequency Atlas of the United States. NOAA Atlas 14 vol. 1.
- Boushaki, F.I., Hsu, K., Park, G.-H., Mahani, S., Sorooshian, S., 2009. Bias adjustment of satellite precipitation estimation using ground measurement: a case study evaluation over the southwestern US. *J. Hydrometeorol.* 10 (5), 1231–1242. <https://doi.org/10.1175/2009JHM1099.1>.
- Burn, D.H., 1990. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resour. Res.* 26 (10), 2257–2265.
- Calheiros, R.V., Zawadzki, I., 1987. Reflectivity-rain rate relationship for radar hydrology in Brazil. *J. Clim. Appl. Meteorol.* 26, 118–132.
- Cavadias, G.S., Ouarda, T.B.M.J., Bobee, B., Girard, C., 2001. A canonical correlation approach to the determination of homogenous regions for regional flood estimation of ungauged basins. *Hydrol. Sci. J.* 46 (4), 499–512.
- Chiang, Y.M., Hsu, K., Chang, F.-J., Hong, Y., Sorooshian, S., 2007. Merging multiple precipitation sources for flash flood forecasting. *J. Hydrol.* 340, 183–196.
- Dalrymple, T., 1960. Flood Frequency Analysis. U.S. Geological Survey, Washington, DC.
- Eldardiry, H., Habib, E., Zhang, Y., 2015. On the use of radar-based quantitative precipitation estimates for precipitation frequency analysis. *J. Hydrol.* 531, 441–453. <https://doi.org/10.1016/j.jhydrol.2015.05.016>.
- Endreny, T.A., Imbeah, N., 2009. Generating robust rainfall intensity-duration-frequency estimates with short-record satellite data. *J. Hydrol.* 371, 182–191. <https://doi.org/10.1016/j.jhydrol.2009.03.027>.
- Gado, T.A., Nguyen, V.T.V., 2016. Regional estimation of floods for ungauged sites using partial duration series and scaling approach. *J. Hydrol. Eng.* [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001439.04016044](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001439.04016044).
- Gourley, J.J., Hong, Y., Flamig, Z.L., Wang, J., Vergara, H., Anagnostou, E.N., 2011. Hydrologic evaluation of rainfall estimates from radar, satellite, gauge, and combinations on Ft. Cobb Basin, Oklahoma. *J. Hydrometeorol.* <https://doi.org/10.1175/2011JHM1287.1>.
- GREHYS (Groupe de Recherche en hydrologie statistique), 1996. Intercomparison of regional flood frequency procedures for Canadian rivers. *J. Hydrol.* 186 (1–4), 85–103.
- Hong, Y., Hsu, K., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteorol.* 43 (12), 1834–1852.
- Hosking, J.R.M., Wallis, J.R., 1993. Some statistics useful in regional frequency analysis. *Water Resour. Res.* 29 (2), 271–281.
- Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-moments. Cambridge University Press, Cambridge, U.K.
- Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. An appraisal of the regional flood frequency procedure in the U.K. flood studies report. *Hydrol. Sci. J.* 30 (1), 85–109.
- Hsu, K., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.* 36. [https://doi.org/10.1175/15200450\(1997\)036<1176:PEFRSI>2.0.CO;2](https://doi.org/10.1175/15200450(1997)036<1176:PEFRSI>2.0.CO;2).
- Hsu, K., Ashouri, H., Braithwaite, D., Sorooshian, S., 2014. Climate Algorithm Theoretical Basis Document (C-ATBD) for Precipitation - PERSIANN-CDR, NOAA NCDC Climate Data Record Program, CDRP-ATBD-0286.
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R. Kidd, C., Nelkin, E.J., Xie, P., 2015. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4.5, p. 26.
- Huffman, G.J., Coauthors, 1997. The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Am. Meteor. Soc.* 78, 5–20.
- Huffman, G.J., Coauthors, 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorology* 8, 28–55.
- Jamandre, C.A., Narisma, G.T., 2013. Spatio-temporal validation of satellite-based rainfall estimates in the Philippines. *Atmos. Res.* 122, 599–608.
- Jiang, S., Ren, L., Yong, B., Yang, X., Shi, L., 2010. Evaluation of high-resolution satellite precipitation products with surface rain gauge observations from Laohahe Basin in northern China. *Water Sci. Eng.* 3 (4), 405–417. <https://doi.org/10.3882/j.issn.1674-2370.2010.04.004>.
- Joyce, R.J., Jonowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* 5, 487–503.
- Katiraei-Boroujerdy, P.-S., Nasrollahi, N., Hsu, K., Sorooshian, S., 2013. Evaluation of satellite based precipitation estimation over Iran. *J. Arid Environ.* 97, 205–219. <https://doi.org/10.1016/j.jaridenv.2013.05.013>.
- Kuligowski, R.J., 2002. A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeorol.* 3, 112–130.
- Lettenmaier, D.P., Potter, K.W., 1985. Testing the flood frequency estimation models using a regional flood generation model. *Water Resour. Res.* 21 (12), 1903–1914.
- Marra, F., Morin, E., Peleg, N., Mei, Y., Anagnostou, E.N., 2017. Intensity-duration-frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern Mediterranean. *Earth Syst. Sci. Hydrol.* <https://doi.org/10.5194/hess-21-2389-2017>.
- Natural Environmental Research Council, 1975. "Flood studies report." Hydrological studies, Vol. 1, London.
- Papalexiou, S.M., Dilynas, Y.G., Grimaldi, S., 2016. Hershfield factor revisited: correcting annual maximum precipitation. *J. Hydrol.* 542, 884–895. <https://doi.org/10.1016/j.jhydrol.2016.09.058>.
- Piman, T., Babel, M.S., Das Gupta, A., Weesakul, S., 2007. Development of a window correlation matching method for improved radar rainfall estimation. *Hydrol. Earth Syst. Sci.* 11, 1361–1372.
- Sorooshian, S., Hsu, K., Braithwaite, D., Ashouri, H., NOAA CDR Program, 2014. NOAA Climate Data Record (CDR) of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR), Version 1 Revision 1. [indicate subset used]. NOAA National Centers for Environmental Information. DOI:10.7289/V51V5BWQ [access date].
- Sorooshian, S., Hsu, K., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates for tropical rainfall. *Bull. Am. Meteorol. Soc.* 81, 2035–2046.
- Su, F., Hong, Y., Lettenmaier, D.P., 2008. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the la plata basin. *J. Hydrometeorol.* <https://doi.org/10.1175/2007JHM944.1>.
- Sveinsson, O.G.B., Salas, J.D., Boes, D.C., 2002. Regional frequency analysis of extreme precipitation in Northern Colorado and Fort Collins flood of 1997. *J. Hydrol. Eng.* 7 (1), 49–63.
- Tasker, G.D., 1982. Comparing methods of hydrologic regionalization. *Water Resour. Bull.* 18 (6), 965–970.
- Ushio, T., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., Kachi, M., Oki, R., Morimoto, T., Kawasaki, Z., 2009. A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Jpn.* 87A, 137–151.
- Wasko, C., Parinussa, R.M., Sharma, A., 2016. A quasi-global assessment of changes in remotely sensed rainfall extremes with temperature. *Geophys. Res. Lett.* 43, 12659–12668. <https://doi.org/10.1002/2016GL071354>.
- Wiltshire, S.E., 1986. Regional flood frequency analysis. I: homogeneity statistics. *Hydrol. Sci. J.* 31 (3), 321–333.
- Wright, D.B., Smith, J.A., Villarini, G., Baek, M.L., 2013. Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. *J. Hydrol.* 488, 150–165. <https://doi.org/10.1016/j.jhydrol.2013.03.003>.
- Xie, P., Arkin, P.A., 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Meteorol. Soc.* 78, 2539–2558.
- Xue, X., Hong, Y., Limaye, A.S., Gourley, J.J., Huffman, G.J., Khan, S.I., Dorji, C., Chen, S., 2013. Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.* 499, 91–99.
- Yilmaz, K.K., Hogue, T.S., Hsu, K., Sorooshian, S., Gupta, H.V., Wagener, T., 2005. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates on hydrologic forecasting. *J. Hydrometeorol.* 6 (4), 497–517.